Synopsis for Registration as a Research Student

for Doctoral Degree

on

Design and Development of Expert System for Prediction of Cardiovascular Disease

by

Ms. Wagre Savita Sambhaji

Research Scholar

Under the guidance of

Dr. Mrs. Kalpana Paithane - Jondhale

Professor and Head

Department of Electronics and Telecommunication Engineering,

Mahatma Gandhi Mission's College of Engineering, Nanded

Research Centre: Department of Computer Science and Engineering,
Mahatma Gandhi Mission's College of Engineering,
Nanded – 431 605, INDIA.
E-mail : mgmcen@mgmcen.ac.in, Phone No. : 02462 222999

Table of Contents

Abstract.....3

1. Introduction.....4

1.1. Classification of cardiovascular disease (CVD) 4

2. Literature Review 5

3. Problem Statement 8

4. Objectives of research 8

5. Research Methodology for expert system 8

5.1 Parameters for Cardiovascular disease (CVD) 9

6. Proposed work plan 10

References.....11

Abstract

Cardiovascular diseases (CVD) are caused by a group of conditions that affect the heart and blood vessels. It is one of the leading causes of death worldwide and responsible for millions of deaths each year. The early and accurate detection of CVD plays a vital role in reducing the risk of disease and its complications. This abstract provides a concise overview of significance, methodologies and applications for prediction of cardiovascular diseases. There is a need to design and develop expert system as a tool for early prediction of cardiovascular diseases that enable proactive interventions, leading to better management and prevention of these diseases. The expert system for cardiovascular disease prediction empowers people and medical professionals to make better clinical decisions, prevent potentially fatal illnesses and treat patients more effectively in the early stage.

1. Introduction

The Cardiovascular disease is one of the largest fields where Artificial intelligence and machine learning tools are used in early prediction and diagnosis. According to estimates, 17.9 million fatalities worldwide in 2019 were attributed to cardiovascular illnesses, or 32% of all deaths. Heart attacks and strokes accounted for 85% of these deaths. High- and middle-class countries account for about 75% of cardiovascular disease-related mortality. The heart or blood arteries are connected to cardiovascular disease. The body's cardiovascular system is in charge of distributing blood throughout the body. The heart, blood capillaries, arteries, and veins make up the cardiovascular system's major components. In the event that any of the cardiovascular system's components are unable to carry out their intended functions, cardiovascular
Illness happens. Instant death may occur from a cardiovascular system failure. The death rates are very high since disease is not diagnosed at an early stage. The expert system help enable us to automatically diagnose cardiovascular diseases in their early stages. The design and development of expert system for cardiovascular disease diagnosis can be used to find minimum disease levels and curing different cardiovascular diseases.

1.1 Classification of Cardiovascular Disease (CVD):

There are various types of chronic diseases such as cardiovascular disease (CVD), a variety of cancers, chronic respiratory illnesses, diabetes are responsible for millions of death every year A collection of disorders affecting the heart and blood arteries is referred to as cardiovascular disease, or CVD. The following are the prevalent forms of cardiovascular disease:

1. Coronary Artery Disease (CAD): This occurs when the blood vessels that supply oxygen and nutrients to the heart muscle become narrowed due to the build-up of plaque (a mixture of fat, cholesterol, and other substances). Heart attacks, angina, and other problems can result from coronary artery disease (CAD).
2. Stroke: A stroke occurs when the blood supply to part of the brain is interrupted or reduced, leading to damage to brain tissue. This can happen either due to a blockage in an artery (ischemic stroke) or bleeding into the brain (hemorrhagic stroke). High blood pressure, high cholesterol, smoking, diabetes, and obesity are risk factors for stroke.

3. Heart Failure: Heart failure occurs when the heart muscle becomes too weak to pump blood effectively. This can result from conditions such as CAD, high blood pressure, or damage to the heart muscle (e.g., from a heart attack). Leg swelling, exhaustion, and shortness of breath are signs of heart failure.

4. Arrhythmias: Too rapid, too slow, or irregular heartbeats are all possible symptoms of arrhythmias, which are irregular cardiac rhythms. Arrhythmias vary in their potential for danger; some are benign, while others may be fatal. Common types of arrhythmias include atrial fibrillation, atrial flutter, and ventricular tachycardia.

A blockage that stops blood from reaching the heart or brain is the primary cause of heart attacks and strokes, which are typically acute events. The most frequent cause of this is an accumulation of fat deposits on the inside walls of the blood arteries supplying the brain and heart. Bleeding from brain vessels or blood clots are two possible causes of strokes.

2. Literature Review

Pavate et al. used Fuzzy logic for the detection of minimum levels of a disease through analysis of the patient's lab tests. They have used expert system and diagnosed the cardiovascular disease which saved both time and the patient's patience. After the systematic diagnosis, the doctor advised the proper medicine to the patient to cure the disease. They have focused on the causes of diabetes and its ability to cause a stroke to affect different diseases in the human body [1]. Khairina et al. investigated the expert system for automatically diagnosing the disease in humans. They have used an automatic fuzzy, rules-based inference engine for the detection of the symptoms of the disease [2].

Machine learning methodologies such as HRFLM (Hybrid Random Forest with linear model) is used. Sunit Mistry used SVM and ANN to forecast cardiovascular disease and fairly reliable in predicting cardiac disease. Comparison of Naïve Bayes, KNN, SVM, Decision Tree, Random Forest with HRFLM is done [3]. M. Gudadhe et al. build an accurate and fast AI based system to predict cardiovascular disease using SVM and ANN[4]. F. S. Alotaibi improved the heart failure prediction accuracy to 90.19% using UCI heart disease dataset [5]. S. Mohan et al. proposed hybrid automatic approaches to diagnose cardiovascular disease by identifying the important features and achieved the accuracy upto 88.07% [6]. S. M. S. Shah et al developed an automatic diagnostic methodology from the heart disease dataset by extracting the key

features using feature selection methods and achieved accuracy 85.12% [7].

A. S. Abedeldjoiaad et al. improved accuracy 92% for heart disease prediction using chi-square feature selection and SVM algorithm in Vietnamese patients [8]. A. S. Abdullah presented data mining model for predicting coronary heart disease in a detailed manner and achieved accuracy in the range of 60%-80%[9]. Chang-Sikson proposed only one algorithm to furnish the final result of heart failure detection and achieved accuracy 88.7%[12]. Jian Ping Li proposed machine learning algorithms such as Naive Bayes, Decision Tree, Random Forest, Logistic Regression, SVM and achieved an accuracy of 92%[14]. Rohit Bharti done comparative analysis among KNN, SVM, Logistic Regression methods and promising results were achieved[15]. Sahoo et al. used only two algorithms and achieved an accuracy of 85.2% and 88.9%[16].

Table 1: Performance Analysis of proposed methods of several authors in cardiovascular disease detection

Sr No	Authors	ML Techniques	Evaluation Measures	Data Set	Summary	Year
1	Dr. Poonam Ghuli[24]	Random Forest	Confusion matrix, Accuracy = 90.16%	This paper uses a total of 76 attributes and 14 sub features.	Only one algorithm has been used to furnish the final result of detecting heart failure detection.	2020
2	Shruti Patil , Mrunal Annadate [18]	KNN, decision tree, linear regression and SVM	Accuracy = 88.9 SVM , K nearest Neighbor=82%	This paper consists of a dataset having 303 records with 14 features.	In comparison to the suggested method, the accuracy gained is lower. It provides 92% accuracy.	2020
3	Chitham baram T,Logesh Kannan N, Gowsalya M[17]	Logistic Regression (LR), Adaptive Boosting (AdaBoostM1), Rule Induction	By majority vote, accuracy is increased by 80.20. LR has the lowest	This paper consists of a dataset having 303 records with 14 features.	In comparison to the suggested method, the accuracy gained is lower. It provides	2020

		Algorithm (FURIA) to build the second model,Multi-Objective Evolutionary Fuzzy Classifier (MOEFC).	accuracy when used without ensemble while AdaBoostM1 has the highest accuracy.		90% accuracy.	
4	Sudareswar Pullela [20]	Random Approaches to oversampling, synthetic minority oversampling, and adaptive synthetic sampling	88.7% accuracy with hybrid random forest. KNN, DT, SVM, and Naive Bayes algorithms were used.	This paper consists of a dataset having 644 records which is not preprocessed. The data set taken from Ten Year CHD.	The dataset was unbalanced	2018
5	Archana Singh, Rakesh Kumar [23]	ANN, DT, Random Forest, SVM, naïve Bayes and KNN.	Accuracy= 90.19%(Decision tree),87.36%(Logistic regression), 89.14%(Random	This paper consists of a dataset having 303 records with 14 features.	The research paper does not contain a detailed explanation on Experimental data analysis. Yet Not containing	2020

From above literature survey, a lot of highlights have been given to the application of machine learning algorithms in the field of cardiovascular diseases. The development of information and communication technology has brought advancement majorly in the field of medical science. Advances in smart hardware have enabled large-scale collection of medical data, and which has led to the advancement in various machine learning algorithms and deep learning methods used in medical science. Researchers have used many new approaches to achieve better results.

The goal of this research work is to create an expert system for the prediction of cardiovascular diseases using artificial intelligence approaches.

3. Problem Statement:

This research is aimed towards design and development of expert system for diagnosis and prediction of cardiovascular disease.

In our research work the following problems can be addressed:

1. Design and develop expert system model for cardiovascular disease diagnosis and prediction that not only provide accurate prediction but also offer insights into the underlying factors contributing to the predicted risk, aiding clinicians in decision-making.
2. Ensure the quality and standardization of data used for training and validating prediction models is crucial.
3. Test and validate the expert system forecasting model for diagnosis of cardiovascular disease.

4. Objectives of Research

RO1. To Design and develop expert system model for cardiovascular disease diagnosis and prediction.

RO2. To propose an efficient algorithm for detection of specific cardiac disorder.

RO3. To use various combinations of ML strategies in terms of keep moving the forecasting model.

RO4. To implement and validate deep learning based algorithm to automatically diagnose and detect cardiovascular disease.

5. Research Methodology for Expert System:

This research is proposing an expert system for medical diagnosis to find accurate anda minimum disease level and present the vital system for heart diseases. Developing an expert system for Cardiovascular Disease (CVD) involves creating a computerized system that can imitate the decision-making abilities of a human expert in the field of cardiovascular health. Such a system can help in early detection, risk assessment, and management of cardiovascular diseases. The methodology involves three stages: (i) creation of a novel classifier using the annotated dataset, (ii) development of a classification model, and (iii) optimization

of the classification model's parameters.

5.1 Parameters for Cardiovascular Disease (CVD)

Cardiovascular disease (CVD) involves a range of risk factors and parameters that healthcare professionals assess to evaluate an individual's cardiovascular health. Here are some key parameters and risk factors associated with CVD:

1. **Blood Pressure:** Hypertension, or high blood pressure, is a significant risk factor for CVD. Blood pressure is measured in millimeters of mercury (mmHg), and normal blood pressure is typically considered to be around 120/80 mmHg.
2. **Cholesterol Levels:** High levels of low-density lipoprotein (LDL) cholesterol, often referred to as "bad" cholesterol, and low levels of high-density lipoprotein (HDL) cholesterol, or "good" cholesterol, are associated with an increased risk of CVD.
3. **Triglyceride Levels:** Elevated levels of triglycerides, a type of fat found in the blood, are also linked to cardiovascular risk.
4. **Blood Glucose Levels:** Elevated blood glucose levels, as seen in diabetes or prediabetes, are associated with an increased risk of CVD.
5. **Body Mass Index (BMI):** BMI is a measure of body fat based on height and weight. One important risk factor for cardiovascular disorders is obesity.
6. **Waist Circumference:** Abdominal obesity, as indicated by an increased waist circumference, is associated with an elevated risk of CVD.
7. **Physical Activity:** Lack of physical activity is a modifiable risk factor for CVD. Regular exercise enhances cardiovascular health overall, lowers blood pressure, and helps people maintain a healthy weight.
8. **Diet**: Cardiovascular disorders can arise as a result of eating a diet heavy in cholesterol, added sugars, saturated and trans fats, and salt. A heart-healthy diet rich in fruits, vegetables, whole grains, and lean proteins is recommended.
9. **Smoking Status:** Tobacco smoke contains chemicals that can damage blood vessels and heart tissue, increasing the risk of CVD. Smoking cessation significantly reduces this risk.
10. **Alcohol Consumption:** Excessive alcohol intake is associated with an increased risk of hypertension and other cardiovascular problems. Moderation is key, and some studies suggest that moderate alcohol consumption may have certain cardiovascular benefits.

11. **Family History:** A family history of cardiovascular diseases can indicate a genetic predisposition. Individuals with a family history may need closer monitoring and preventive measures.
12. **Age and Gender:** Advancing age is a non-modifiable risk factor for CVD, and men generally have a higher risk than premenopausal women. However, the risk increases in postmenopausal women.
13. **Stress Levels:** Chronic stress can contribute to CVD risk. Stress management techniques, such as meditation or exercise, may be beneficial. It is found that these parameters are interconnected, and the overall cardiovascular risk assessment considers multiple factors. Regular health check-ups, lifestyle modifications, and early intervention can help manage these risk factors and reduce the likelihood of developing cardiovascular diseases.

6. Proposed Work Plan

Table 2: Year-wise Plan of Work and target to be achieved

Year	Work	Research Objectives	Target to be achieved
I (2023-24)	Literature Review, Tools understanding for problem solving, Learning new technologies, implementation of algorithms, Result analysis and Research article writing.	RO1	Journal Paper I
II (2024-25)	Working on detection of cardiac disorders, Working on data pre-processing on dataset, Design and implement an algorithm, Result analysis and Research article writing.	RO2 RO3	Conference Paper I Journal Paper II
III (2025-26)	Pre Ph.D. Synopsis Submission, Thesiswriting, Thesis Submission	RO4	Conference Paper II /Book Chapter I

References

[1] Pavate, A., Nerurkar, P. et al., “Early Prediction of Five Major Complications Ascends in Diabetes Mellitus Using Fuzzy Logic”. In Soft Computing in Data Analytics; Springer: Berlin/Heidelberg, Germany, 2019; pp. 759–768.

[2] Khairina et al., “Automation Diagnosis of Skin Disease in Humans using Dempster-Shafer Method ” E3S Web Conf. 2018, 31, 11006.

[3] Sunit Mistry, “Efficient Prediction of Heart Disease using Cross Machine Learning Techniques”, IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers(IPEC), 2022.

[4] M Gudadhe et al., “Decision support system for heart disease based on support vector machine and artificial neural network”, in Proc. Int. Conf. Comput, Commun, Techno,(ICCCT), Sep 2010.

[5] F. S. Alotaibi, “Implementation of machine learning model to predict heart failure disease”, published in Science Direct, September 2012.

[6] S. Mohan et al., “Effective heart disease prediction using hybrid machine learning techniques”, IEEE Access, vol. 7, 2019.

[7] S. M. S. Shah et al., “Support vector machine-based heart disease diagnosis using feature subset, wrapping selection and extraction methods”, Comput. Electr. Eng., vol 84, June 2020.

[8] F. Z. Abedeldjoiaad et al., “A hybrid Approach for Heart Disease Diagnosis and Prediction using Machine Learning Techniques”, in Proceedings Int. Conf. Comput. Sci. Eng. (UBMK), Antalya, Turkey, August 2020.

[9] A. S. Abdullah and R. R. Rajalaxmi, ‘‘A data mining model for predicting the coronary heart disease using random forest classifier,’’ in Proceedings Int. Conf. Recent Trends Comput. Methods, Communication. Controls, Apr. 2012.

[10] Senthilkumar Mohan Chandrasegar Thirumalai and Gautam Srivastavar, ‘‘Effective Heart Diesease using Hybrid Machine Learning models,’’in Proc. IEEE 1st Int. Conf. Control, Meas. Instrum. (CMI), Jan. 2016.

[11] Abdulwahab Ali Al Mazroei,"Survival prediction among heart patients using machine learning techniques", published in Aims press conference, pp. 22-24, Nov,2021.

[12] Chang-SikSon, "Decision-making model for early diagnosis of congestive heart failure using rough set and decision tree approaches", published in Science Direct,pp.

12–13,Sept 2012.

[13] Fahd Saleh Alotaibi,"Implementation of Machine Learning Model to Predict Heart Failure Disease",published in the Semantic Scholar website ,pp.11–16,March 2019.

[14] Jian Ping Li,"A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms", published in Hindawi conference. 24-25,Dec 2018.

[15]Rohit Bharti, "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning," published a journal in Hindawi, vol. 20, no. 1, pp. 86–94 April 2021.

[16]Sahoo, Prasanta Kumar and Jeripothula, Pravalika, Heart Failure Prediction Using Machine Learning Techniques (December 15, 2020). Available at SSRN: https://ssrn.com/abstract=3759562 or http://dx.doi.org/10.2139/ssrn.

[17]Pouriyeh, S., Vahid, S., Sannino, G., Pietro, G. D., Arabnia, H., Gutierrez, J., "A comprehensive investigation and comparison of machine learning techniques," IEEE Symposium on Computers and Communications (ISCC),pp. 422–432, Dec 2017.

[18]Shruti Patil, Mrunal Annadate ," Implementation of Machine Learning Model to Predict Heart Problem," published in IRJTE organization ,pp- 14-15 ,July 2022.

[19]Rahul Katarya; Polipireddy Srinivas, "Predicting Heart Disease at Early Stages using Machine Learning: A Survey, A literature review," Eur. J. Oper. Res., vol. 189, no. 1, pp. 194–207, September 2020.

[20] Sundareshwar Pullela, "Machine Learning Techniques For Heart Disease Prediction", Proceedings Internet Technol. Appl. (ITA), Sep. 2017, pp. 110–115, 2020.

[21] Amin Ul Haq, Jian Ping Li, Muhammad Hammad Memon, Shah Nazir, and Ruinan Sun, "A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms," published a journal in Hindawi ,vol. 20, no. 1, pp. 88- 97,June 2018

[22] Fatma Zahra Abdeldjouad, Menaouer Brahami, Nada Matta, "A Hybrid Approach for Heart Disease Diagnosis and Prediction Using Machine Learning Techniques," in Proceedings Int. Conf. Comput. Sci. Eng. (UBMK), Antalya, Turkey,Oct. 2017, pp. 719–723,August 2020

[23] Archana Singh, "Heart Disease Prediction Using Machine Learning Algorithms," Published in the Semantic Scholar website ,pp.43-46, Nov 2020.

[24] Dr Poonam Ghuli,J. P. Kelwade, S. S. Salankar, “Radial basis function neural network for prediction of cardiac arrhythmias based on heart rate time series,’’ in Proceedings IEEE 1st Int. Conf. Control, Meas. Instrum. (CMI), Jan. 2016,pp. 454–458.

Ms. Wagre Savita Sambhaji
Research Scholar

Dr. Kalpana Paithane-Jondhale
Ph.D. Supervisor